

В.М. ТАРАН, ст. викл., Європейський університет, Ялтинська філія (Ялта)

ПРАКТИЧНЕ ВПРОВАДЖЕННЯ РОЗРОБЛЕНИХ МЕТОДІВ ПРОГНОЗУВАННЯ ЗСУВНИХ ПРОЦЕСІВ ПІВДЕННОГО БЕРЕГА КРИМУ

Розглянуто методи регресійного аналізу прогнозування зсувних процесів Південного берега Криму для довгострокового прогнозу та мережі довіри Байєса для короткострокового, а також побудовано відповідні математичні моделі. На основі розроблених методів побудована програмна система, що дозволяє аналізувати та прогнозувати катастрофічні наслідки цих процесів при допомозі даних багаторічних спостережень за кліматичними та гідрогеологічними факторами. Наведено концептуальну модель системи та її технічний опис. Система має гібку архітектуру, що дозволяє її доповнювати додатковими модулями. Іл.: 6. Табл.1. Бібліогр.: 10 назв.

Ключові слова: прогнозування зсувних процесів Південного берега Криму, регресійний аналіз, мережі довіри Байєса.

Постановка проблеми. Моніторинг, моделювання та прогнозування зсувних процесів Південного берега Криму є складною задачею аналізу і прогнозування динамічних процесів довільної природи, що ускладнюється присутністю в даній системі чинника ризику непередбачених ситуацій [1]. В цьому регіоні існують організації, що ведуть спостереження за кліматичними, екологічними та геологічними процесами, які обладнані новітніми технічними пристроями, але ці дані ніяк не пов'язані між собою, тобто відсутнє єдине технічне бюро обробки набутої інформації та вироблення оптимального рішення по впровадженню протизсувних заходів [2].

Аналіз літератури. У наукових дослідженнях [3 – 6], що розглядають прогнозування зсувних процесів Південного берега Криму основна увага приділяється геологічному та геоморфологічному дослідженню схилів та ділянок, на яких відбулися руйнівні зсуви, а саме прогнозування базується більшою мірою на нескладних математичних моделях або лише на інтуїції фахівця. Таким чином, в цих дослідженнях обмеженим є використання сучасних інформаційних технологій, а також приділяється недостатньо уваги задачам інформаційної підтримки та автоматизації планування й управління протизсувними заходами в нестабільних умовах протікання цих процесів.

На сучасному етапі накопичено різномірні дані багаторічних кліматичних, геологічних, екологічних, астрономічних спостережень, але науковці певної галузі при моделюванні використовують лише ті дані, що стосуються безпосередньо цієї галузі, і ті моделі, які себе вже виправдали в попередніх дослідженнях. Таким чином, накопичена інформація використовується не в повному обсязі, не системно, не використовуючи сучасні інформаційні технології.

Мета статті – системний аналіз динаміки зсувних процесів та розробка інтелектуальної системи, яка спрощуватиме процес вироблення і прийняття управлінських рішень.

Використання методів прогнозування часових рядів не завжди забезпечує потрібну точність результатів обробки інформації, тому необхідно розвивати технологію оцінювання і прогнозування головних показників, динаміки внутрішніх та зовнішніх факторів розвитку процесу з метою покращення точності короткострокового та довгострокового прогнозування. Отже, актуальною стає проблема створення системи підтримки прийняття рішень (СППР), яка б допомагала приймати ефективні управлінські рішення та використовувала не один, а декілька методів прогнозування і на їх основі будувала би узагальнений прогноз.

Основні моделі, що використовуються в інтелектуальній системі при прогнозуванні зсувних процесів Південного берега Криму. Південний берег Криму є унікальним регіоном України не лише завдяки географічному розташуванню і клімату, а ще й тому, що на цій достатньо вузькій полосі узбережжя розміщується велика кількість різномірних станцій спостереження. Це й метеорологічні станції (в смт. Нікіті, в Лівадії, в Місхорі, на горі Ай-Петрі); сейсмічна станція, яка має новітнє обладнання; обсерваторії (в Кацівелі, в Сімеїзі); протизсувне управління і Ялтинська гідрогеологічна станція, що безпосередньо формує базу даних стосовно зсувів; також в Кацівелі є філіал Севастопольського гідрогеологічного інституту моря, який має власні станції спостереження за кліматичними, морськими та гідрогеологічними змінами. Проте, вся зібрана інформація підпорядковується різним галузям наук, тому обробляється за допомогою відомих в цій галузі математичних моделей і передається до Києва у відповідні установи та відомства, а на місці не проводяться комплексні розрахунки, в єдиній системі, враховуючи всі фактори ризику і прогножуючи їх катастрофічні наслідки.

Наведемо деякі моделі та методи, що описують зсувні процеси Південного берега Криму: однофакторна регресійна модель, авторегресійна, комплексна множинна лагова авторегресійна модель, метод групового врахування аргументів та метод комплексування аналогів.

По-перше, розглянемо чинники, що сприяють активізації зсувних процесів: опаді, сейсмічна активність та сонячна активність. Новизною в даному дослідженні є те, що будемо пов'язувати між собою чинники попереднього року і результируючий показник наступного, тобто кожна модель буде лаговою (з лагом рівним 1 року). Опаді враховувалися багатьма іншими науковцями, але такі прогнози мали невелику достовірність. Отже, для їх покращення будемо накопичувати кількість опадів за два роки (O). В якості фактора сейсмічної активності візьмемо кількість поштовхів впродовж року з магнітудою, що перевищує 8,5 K, які відбулися на відстані до 200 км від Південного берега Криму (CeA). Сонячну активність будемо вимірювати за

річною кількістю чисел Вольфа (CoA). Зсувний процес опишемо за допомогою кількості активних зсувів за рік (A3). Розглянемо наступні моделі:

однофакторну регресійну модель (P):

$$A3(t) = a_0^1 + a_1^1 O(t-1);$$

авторегресійну модель (AP):

$$A3(t) = a_0^2 + a_1^2 A3(t-1);$$

комплексну множинну лагову авторегресійну модель (ЛАР):

$$A3(t) = a_0^3 + a_1^3 O(t-1) + a_2^3 CeA(t-1) + a_3^3 CoA(t-1) + a_4^3 A3(t-1) + a_5^3 (t-1)$$

(цю модель доповнено авторегресійною складовою $A3(t-1)$ та значенням попереднього року $(t-1)$);

модель методу групового врахування аргументів (МГВА):

$$Y = a_0^4 + \sum_{i=1}^N a_i^4 x_i + \sum_{j=1}^N \sum_{i=1}^N a_{ij}^4 x_i x_j + \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N a_{ijk}^4 x_i x_j x_k + \dots;$$

модель методу комплексування аналогів (КА):

$$E_i = \sqrt{\sum_{j=1}^n (x_{ij} - x_j)^2} \rightarrow \min, \quad i = \overline{1, k},$$

де $a_0^1, a_1^i, (i = \overline{1, 3}), a_2^3, a_3^3, a_4^3, a_5^3, a_0^4, a_i^4, (i = \overline{1, N}), a_{ij}^4, (j, i = \overline{1, N}, i \leq j), a_{ijk}^4, (i, j, k = \overline{1, N}, j \leq i, k \leq j)$ – постійні коефіцієнти.

Позначимо E_{a1} – найменшу відстань між факторами, E_{a2} – ту, що слідує за найменшою, y_{a1} – величина, найближча до прогнозованої невідомої величини y , а y_{a2} – друга за нею по відстані, тоді результат визначаємо за формулою

$$y = \frac{y_{a1} \cdot E_{a2} + y_{a2} \cdot E_{a1}}{E_{a1} + E_{a2}}.$$

Наведемо критерії вибору кращих моделей з множини оцінюваних кандидатів:

– коефіцієнт множинної детермінації

$$R^2 = 1 - \frac{\sum_{i=1}^n (\tilde{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

де y_i – значення показника для $i=1, 2, \dots, n$; \tilde{y}_i – прогнозна величина;

– середній квадрат похибки

$$e^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2 ;$$

– критерій Дарбіна-Уотсона для перевірки на корельованість похибок моделі

$$DW = 2 - 2\rho ,$$

де ρ коефіцієнт корельованості похибок моделі між собою;

- t -статистика Стюдента для перевірки значимості параметрів моделі;
- інформаційний критерій Акайке

$$AIC = n \ln \sum_{i=1}^n e_i^2 + 2k ,$$

де n – довжина вибірки даних; k – число параметрів моделі;

- критерій Байєса-Шварца

$$ABC = n \ln \sum_{i=1}^n e_i^2 + k \ln(n) ;$$

- F -критерій Фішера для перевірки адекватності моделі в цілому

$$F = \frac{(n-m-1) \sum_{i=1}^n (\tilde{y}_i - \bar{y})^2}{m \sum_{i=1}^n (y_i - \tilde{y})^2} ,$$

де n – кількість спостережень; m – кількість змінних.

Всі відомі показники, що використовуються для оцінки якості отриманого прогнозу можна розділити на три групи: абсолютні, порівняльні та якісні. До абсолютних показників якості прогнозів відносяться такі, що дозволяють кількісно оцінити величину похибки прогнозу в одиницях вимірювання прогнозованого об'єкту або у відсотках. В практиці частіше використовують такі формальні статистики перевірки якості прогнозу похибки точності:

Середня абсолютна похибка MAE (mean absolute error):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_i| ,$$

де: y_i – значення періоду, \tilde{y}_i – прогнозна величина для $i = 1, 2, \dots, n$.

Середня абсолютна відсоткова похибка MAPE (mean absolute percentage error):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \tilde{y}_i|}{|y_i|} \cdot 100\% .$$

Середньоквадратична похибка RMSE (root mean squared error):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y})^2} .$$

Коефіцієнт нерівності Тейла U :

$$U = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i)^2} + \sqrt{\frac{1}{n} \sum_{i=1}^n (\tilde{y}_i)^2}}.$$

В таблиці можемо для порівняння навести результати оцінювання запропонованих моделей за наведеними критеріями.

Таблиця

Оцінювання запропонованих моделей

Регрес. Статист.	Регресія	Авто- регресія	ЛАР (5 факторів)	МГВА (15 факторів)	Комплексування аналогів
Множ. R	0,641	0,708	0,803	0,896	0,897
R^2	0,411	0,501	0,645	0,802	0,804
Среднекв.откл.	45,42	41,80	37,18	32,67	45,97
e^2	82504,85	69889,73	49753,73	27746,62	84889,07
DW	0,912	2,02	1,852	2,28	1,587
F	27,96	40,20	13,08	7,03	7,14
RMSE	44,32	58,02	35,82	44,33	45,50
MAE	32,92	45,32	30,19	33,86	31,06
MAPE	35,30%	51,71%	31,39%	36,49%	26,8%
U	0,154	0,203	0,124	0,144	0,165

Таким чином, кращими моделями можна вважати третю і четверту, тобто комплексну множинну лагову авторегресійну модель і модель, отриману за допомогою методу групового врахування аргументів. На третьому місці – модель, яка отримана за допомогою методу комплексування аналогів.

Для врахування всіх результатів моделювання побудуємо узагальнений прогноз, який містить всі оцінки моделювання, але не є їх середнім арифметичним тому, що ці моделі не є рівноцінними. Виберемо в якості вагових коефіцієнтів R^2 – коефіцієнти детермінації кожної з моделей. Отже, узагальнений результат моделювання будемо розраховувати за формулою:

$$A3_{\text{узаг.}} = \frac{A3_1 \cdot R_1^2 + A3_2 \cdot R_2^2 + A3_3 \cdot R_3^2 + A3_4 \cdot R_4^2 + A3_5 \cdot R_5^2}{R_1^2 + R_2^2 + R_3^2 + R_4^2 + R_5^2}.$$

Такий прогноз будемо вважати довгостроковим тому, що мінімальний інтервал прогнозування 1 рік. Для отримання прогнозу на менший термін вперед (тиждень, місяць та ін.) побудуємо мережу довіри Байєса.

Мережа Байєса (МБ) представляє орієнтований ациклічний граф, вершини якого є дискретні випадкові змінні зі скінченим числом станів, а

ребра є причинними зв'язками між ними і характеризуються таблицею безумовних ймовірностей переходів з одного стану до іншого під впливом збурень. Отже, МБ характеризується парою чисел (G, P) , де $G = \langle X, E \rangle$ – спрямований ациклічний граф на скінченій множині X , елементи якої поєднані між собою сукупністю орієнтованих ребер E , а P – множина умовних розподілів ймовірностей [2, 8].

На першому рівні розташовано фактори впливу (опаді, сейсмічна активність, сонячна активність, кількість активних зсувів, що мали місце в попередньому терміні) та управляючий фактор (У – укладені кошти в укріплення зсувонебезпечних ділянок). На другому рівні – показник кількості активних зсувів, на третьому – показник ризику часу, впродовж якого можуть статися катастрофічні зсуви, на четвертому – узагальнений показник обсягу витрат на укріплення схилів та подолання руйнівних наслідків.

За допомогою байєсівських мереж формується логічний висновок відносно стану цільових вузлів ймовірнісної моделі об'єкта на основі значень спостережень за параметрами, що контролюються (значення вузлів пасивного контролю) і результатів додаткового тестування (значення вузлів активного контролю). Причому в якості значень вузлів пасивного контролю можуть використовуватися як поточні спостереження, так і результати прогнозування спостережень, що реалізуються блоком оцінки і прогнозування стану процесу (див. рис.1).

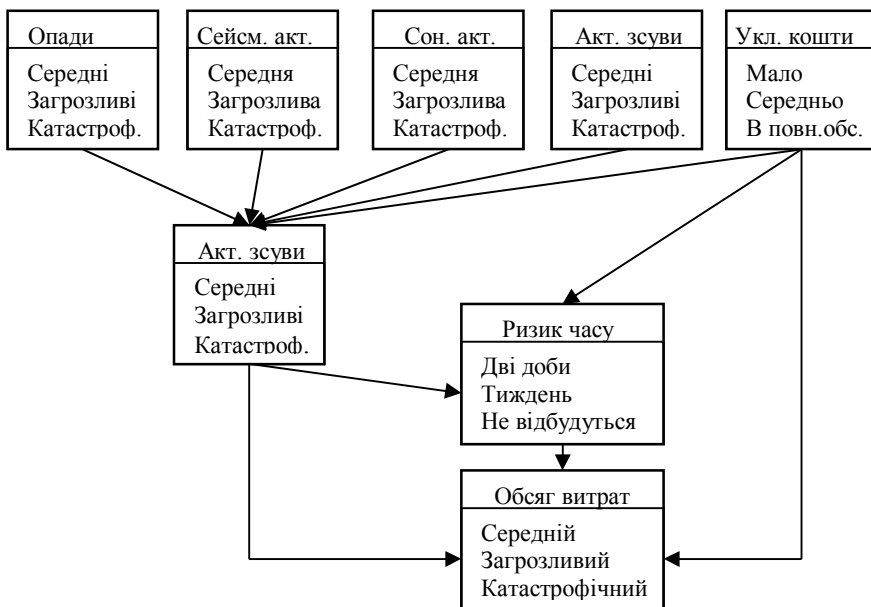


Рис.1. Мережа довіри Байєса для короткострокового прогнозування зсувних процесів Південного берега Криму

СППР при прогнозуванні зсувних процесів Південного берега Криму.

Для прогнозування зсувних процесів Південного берега Криму запропонована СППР, яка має концептуальну структуру, що показано на рис. 2.

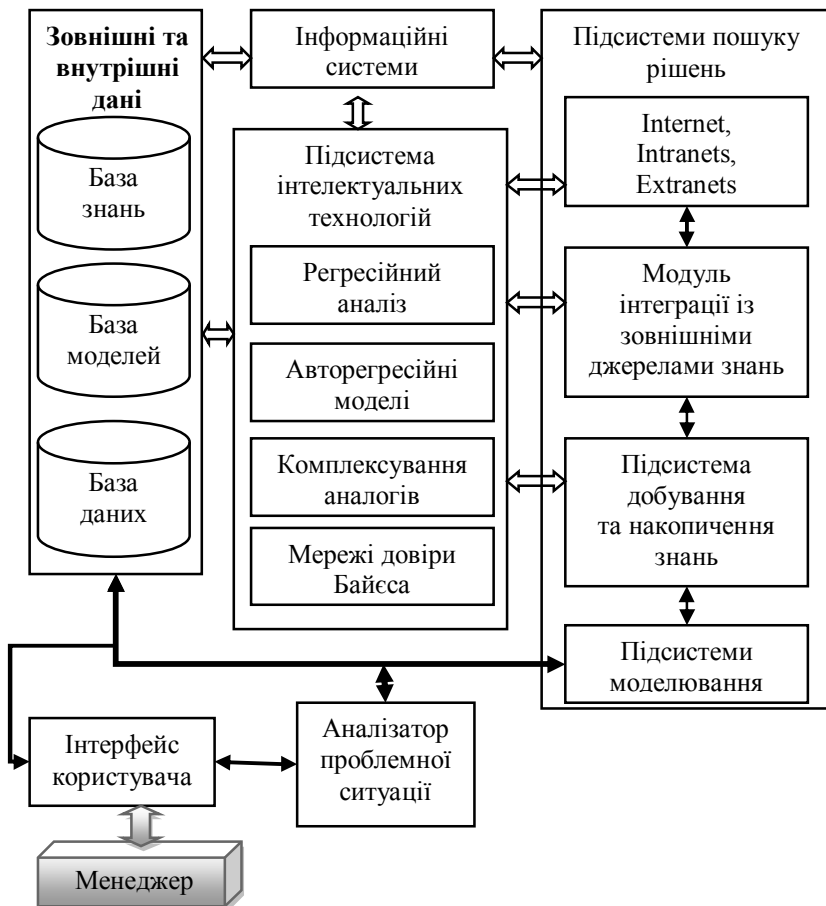


Рис. 2. Концептуальна структура інтелектуальної системи підтримки прийняття рішень

При вивченні та аналізі даних, що пов'язані з зсувними процесами Південного берега Криму, виявилось, що найбільш впливові теоретично чинники кількісно пов'язані дуже слабо. Спостереження за кліматичними, гідрогеологічними та метеорологічними даними ведуться різними організаціями і містять набір результатів спостережень майже за п'ятдесят років, а по деяким параметрам перевищує одне століття, але ці організації

фактично не пов'язані між собою оскільки відносяться до різних галузей, а значить відсутній обмін даними поміж ними для встановлення певних попереджуючих факторів.

В даному дослідженні пропонується система побудови прогнозів з використанням декількох методів прогнозування, якість та точність яких оцінюється експертною системою, що є складовою модуля LongTimeForecast. Порівняння отриманих довгострокових прогнозів проводиться на основі рішення задач за допомогою традиційних методів прогнозування: регресійного аналізу, авторегресії, кореляційного аналізу, комплексування аналогів; а короткострокового – мереж довіри Байєса [8, 10]. При оцінюванні системою отриманих прогнозів вибирається найкращий (оптимальний) або будується узагальнений прогноз, котрий є комбінацією прогнозів, побудованих за допомогою різних методів.

Модуль LongTimeForecast містить необхідні інструменти та процедури для рішення таких задач, як:

- попередня обробка та аналіз даних, необхідних для розробки прогнозів;
- вибір методу прогнозування,
- оцінка якості прийнятої прогнозної моделі та метода.

При функціонуванні СППР проводиться аналіз якості обраної моделі та аналіз і оцінювання прогнозу (рис. 3). Якщо якість моделі відповідає обраному критерію, то будується прогноз, який в свою чергу перевіряється за допомогою обраних характеристик. Якщо прогноз не відповідає певному критерію, то отримані моделі та результати прогнозування відкидаються і моделювання починається знову, спираючись на оновлені дані.

Таким чином, питання про якість прогнозу зводиться до визначення допустимої середньої похибки, відносної похибки у відсотках, коефіцієнту кореляції, критерію Дарбіна-Уотсона та коефіцієнта нерівності Тейла U . В дослідженнях здійснювалося порівняння прогнозів, отриманих за допомогою мереж довіри Байєса з прогнозами, отриманими з використанням класичних статистичних методів.

При отриманні прогнозу побудована система створює графічне відображення – множину форматів графічного відображення для візуального сприйняття таблиць і графіків [8 – 10]. Для короткострокового прогнозу використовуються мережі Байєса, що надають результат у вигляді ймовірності появи даного розвитку подій.

За допомогою вибору варіанта протікання даного процесу фахівець відповідає на питання стосовно поточного стану чинників, а система моделює прогноз в залежності від обраного варіанту результату.

В залежності від обраного варіанту прогнозу результатами короткострокового прогнозування можуть бути:

- кількість активних зсувів (незначна, загрозна – позаштатна ситуація, катастрофічна);

- ресурс часу, впродовж якого має відбуватися зсувний процес (найближчі дві доби, найближчий тиждень, або зсуви не відбудуться зовсім);
- обсяг коштів на відновлення ушкоджених об'єктів (невеликий, помірний, катастрофічний).

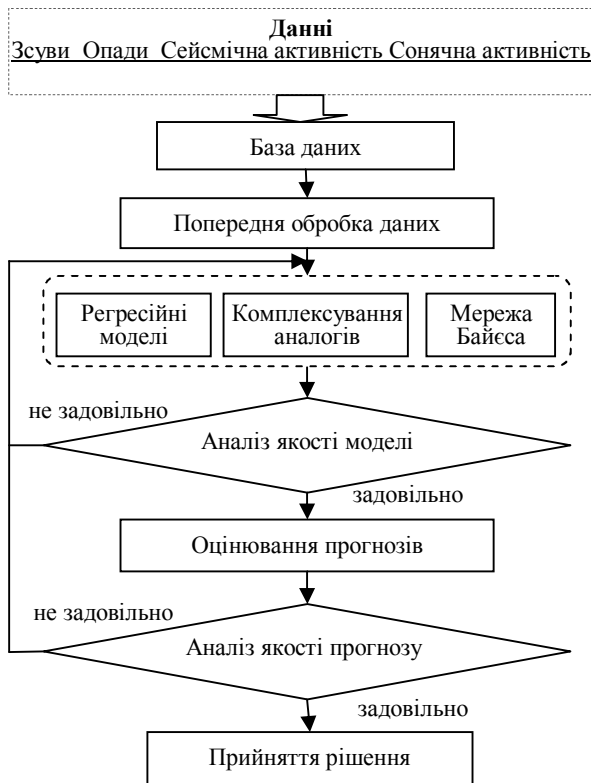


Рис. 3. Структура СППР

Довгострокові прогнозування реалізують п'ять методів: регресійний аналіз, авторегресія, лагова авторегресія, комплексування аналогів та метод групового врахування аргументів. Для оцінки кількості зсувів в наступному році (можна обирати на рік вперед, на два, на п'ять) треба ввести значення кількості опадів за два попередніх роки, кількість поштовхів за попередній рік, що перевищують магнітуду 8 К і не перевищують відстань від епіцентру 200 км, а також число Вольфа (сонячна активність) і кількість активних зсувів в попередньому році.

Всі вхідні данні вводяться за попередній рік, щоб отримати прогноз на наступний (такий вибір моделі, що пов'язує данні попереднього і наступного

років, запропоновано вперше). Показники всіх факторів підсумовуються за рік (хоча є данні за більш короткий термін, ніж рік) тому, що результуючий показник – кількість зсувів фіксується за цілий рік, тобто раз на рік фахівці Ялтинської гідрогеологічної партії, яка відповідає за результати обстеження вже відомих і нових зсувів, виїжджають на місце і фіксують наявність активності того чи іншого зсуву, описують, вимірюють та заносять в Кадастр зсувів Криму. В разі появи нових зсувів приходить повідомлення від відповідних служб і фахівці аналогічно все фіксують (але знову ж таки раз на рік). Таким чином, ці моделі розглядаємо як довгострокові, тобто прогноз оцінюється не менше ніж на рік вперед.

Отже, при використанні СПІР та створенні єдиного технічного бюро обробки інформації та прогнозування можливо отримувати достовірні оцінки прогнозів зсувних процесів Південного берега Криму та оптимізувати витрати (за допомогою мережі довіри Байеса) на впровадження протизсувних заходів і подолання катастрофічних наслідків цих процесів.

Висновки. Запропоновано поєднати методи регресійного аналізу, комплексування аналогів та мережі довіри Байеса для моделювання і прогнозування зсувних процесів Південного берега Криму, які використовують попередні відомі данні для отримання оцінок прогнозу в майбутньому, що є новим в моделюванні цих процесів. Таким чином, для прогнозування запропоновано п'ять математичних моделей (однофакторна регресія, авторегресія, комплексна множинна лагова авторегресія, моделі методу групового врахування аргументів та метода комплексування аналогів) і критерії оцінки моделей та оцінки прогнозу. Результати розрахунків наведено у вигляді таблиці.

Побудовано концептуальну структуру інтелектуальної системи підтримки прийняття рішень відносно зсувних процесів Південного берега Криму, виділено її головні підсистеми.

Реалізовано інтелектуальну систему у вигляді програмного продукту на мові програмування Delphi2009.

В подальшому передбачається при виводі результату поєднувати всі отримані висновки в один, а також показувати ефективність прогнозу за допомогою не тільки двох обраних критеріїв, а додати в розрахунки й інші розглянуті критерії.

Список літератури: 1. *Селін Ю.М.* Системний аналіз екологічно небезпечних процесів різної природи / *Ю.М. Селін* // Системні дослідження та інформаційні технології. – 2007. – № 2. – С. 22 – 32. 2. *Таран В.М.* Інтелектуальні системи моделювання і прогнозування зсувних процесів Південного берега Криму з використанням мережі довіри Байеса / *В.М. Таран* // Актуальні проблеми економіки. – 2008. – № 10. – 198 – 206 с. 3. *Юдин В.В.* Геологія Криму на основі геодинAMI / *В.В. Юдин*. – С-ктйвкар, 2000. – 44 с. 4. *Трофимов А.М.* Математическое моделирование в геоморфологии склонов / *А.М. Трофимов, В.М. Московкин*. – Издательство Казанского университета, 1983. – 220 с. 5. *Круцик М.Д.* Захист гірських автомобільних доріг від зсувів / *М.Д. Круцик*. – "Коломия". – 2003. – 425 с. 6. *Логвинова К.Т.* Клімат и опасные гидрометеорологические явления Крыма / *К.Т. Логвинова, М.Б. Барабаш*. – Л.: "Гидрометеоздат", 1982.

– 320 с. 7. *Малеева О.В.* Методологические основы системного анализа качества проектов и программ развития производства: Автореферат диссертации на соискание научной степени доктора технических наук. / *О.В. Малеева.* – Х.: Нац. авиац. ин-т "ХАИ", 2003. – 34 с. 8. *Терентьев А.Н.* Использование байесовских сетей в кредитном скоринге / *А.Н. Терентьев, П.И. Бидюк* // Збірка наукових праць міжнародної наукової конференції ISDMC'2008 у трьох томах. Т3 (частина 2). – С. 101 – 103. 9. *Таран В.М.* Навчання мережі Байеса при моделюванні зсувних процесів Південного берега Криму / *В.М. Таран* // Искусственный интеллект. – 2008. – № 3. – С. 600 – 610. 10. *Терентьев А.Н.* Байесовская сеть – инструмент интеллектуального анализа данных / *А.Н. Терентьев, П.И. Бидюк, Л.А. Коршевнюк* // Проблемы управления и информатики. – 2007. – № 4. – С. 83 – 92.

Статья представлена д.т.н. Бидюк П.И.

УДК 519.876:55.435.62

Практическое внедрение разработанных методов прогнозирования оползневых процессов Южного берега Крыма / Таран В.Н. // Вестник НТУ "ХПИ". Тематический выпуск: Информатика и моделирование. – Харьков: НТУ "ХПИ". – 2010. – № 21. – С. 162 – 172.

Рассмотрены методы регрессионного анализа прогнозирования оползневых процессов Южного берега Крыма для долгосрочного прогноза и сети доверия Байеса для краткосрочного, а также построены соответствующие математические модели. На основе предложенных методов построена программная система, позволяющая анализировать и прогнозировать катастрофические последствия этих процессов при помощи данных многолетних наблюдений за климатическими и гидрогеологией факторами. Приведена концептуальная модель системы. Система имеет гибкую архитектуру, что позволяет ее расширять дополнительными модулями. Ил.: 3. Табл.: 1. Библиогр.: 10 назв.

Ключевые слова: прогнозирование оползневых процессов Южного берега Крыма, регрессионный анализ, сети доверия Байеса.

UDC 519.876:55.435.62

Practical introduction of the developed methods of prognostication of landslide processes of the South bank of Crimea / Taran V.N. // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkiv: NTU "KhPI". – 2010. – № 21. – P. 162 – 172.

There are considered regressive analysis methods of prognostication of landslide processes of the Southern Coast of Crimea for a long-term prognosis and Bayesian trust networks for a short-term prognosis. The programmatic system is built on the basis of the developed methods, allowing to analyses and forecast the catastrophic consequences of these processes through these long-term looking after climatic and by a geohydrology by factors. Conceptual model of the system and its technical description, is resulted. The system has flexible architecture, that allows it to extend the additional modules. Figs: 3. Tabl. 1. Refs: 10 titles.

Key words: prognostication of landslide processes of the Southern Coast of Crimea, regressive analysis, Bayesian trust networks.

Поступила в редакцию 29.09.2009